

Smart Lake Rejuvenation with AI-Powered Environmental Monitoring

Sindhu T. S¹, Ms. Akshatha Reddy²

^{1,2}Department of studies in Computer Science (M.Sc.), Davangere University, Shivangotri, Davangere-577007, Karnataka, India.

ABSTRACT

Freshwater pollution has become a major environmental problem that affects both ecosystems and public health. Traditional water quality monitoring methods are manual, infrequent, and unable to provide real-time information. This project, Smart Lake Rejuvenation with AI-Powered Environmental Monitoring, presents a simulation-based system for continuous lake water quality monitoring. The system measures key parameters such as pH, temperature, turbidity, and dissolved oxygen. It uses the Random Forest algorithm to classify water quality and the Isolation Forest algorithm to detect anomalies. Simulated sensor data is used to demonstrate real-world monitoring without physical IoT devices. The system is developed using FastAPI, React.js, and SQLite. It provides a user-friendly dashboard that displays live readings, lake health score, charts, pollution alerts, and future risk predictions. The project also identifies possible pollution sources and explains AI-based results in a simple way. This Digital Twin prototype offers a cost-effective solution for smart lake monitoring and can be extended with real IoT sensors and cloud technology in the future to support sustainable lake management.

Keywords: Water quality monitoring, Machine Learning, Random Forest, Isolation Forest, Anomaly Detection, Explainable AI, Digital Twin, Smart Lake, IoT, Environmental Monitoring.

I. INTRODUCTION

Lakes and reservoirs are important sources of drinking water, biodiversity, groundwater recharge, and livelihoods. However, rapid urbanization, industrial growth, and agricultural activities have increased water pollution through sewage, industrial waste, and chemicals. This has caused serious problems such as eutrophication, algal blooms, fish deaths, and poor water quality. Traditional lake monitoring relies on manual sampling and laboratory testing, which is time-consuming, expensive, and unable to provide continuous monitoring. As a result, pollution is often detected only after significant environmental damage has occurred. Recent advancements in the Internet of Things (IoT), Artificial Intelligence (AI), and cloud computing have made real-time water quality monitoring possible. This project, Smart Lake Rejuvenation with AI-Powered Environmental Monitoring, uses AI and simulated sensor data to monitor important water quality parameters, detect pollution, predict future risks, and provide timely alerts. The system supports proactive lake management and can be upgraded with real IoT sensors for sustainable environmental monitoring.

PROBLEM STATEMENT:

Despite environmental protection laws and the National Lake Conservation Plan, many lakes continue to face severe pollution due to the lack of timely and continuous monitoring. Traditional water quality assessment is manual, infrequent, and reactive, causing delays in identifying pollution events. Existing systems cannot predict future water quality changes or provide clear explanations for pollution sources. Continuous monitoring solutions are often expensive, limiting their large-scale deployment. In addition, there is no integrated platform that combines data collection, AI-based analysis, visualization, and early warning alerts. This project, SLIM-AI Lite, addresses these challenges by providing a cost-effective, AI-powered Digital Twin prototype that continuously monitors water quality, detects anomalies, predicts pollution risks, explains possible causes, and displays results through an interactive dashboard for effective lake management.

OBJECTIVES:

The following official project objectives, adopted without modification as prescribed by , guide the design and implementation of SLIM-AI Lite:

1. To continuously monitor water quality parameters such as pH, turbidity, temperature, and dissolved oxygen.
2. To detect pollutant sources and contamination levels using AI-based models.
3. To predict future water quality trends and algae bloom risks.
4. To automate alerts for government bodies and environmental agencies.
5. To enhance aquatic ecosystem health and restore biodiversity.
1. To promote cost-effective and sustainable lake management strategies.

II. RELATED WORK:

1. IoT-Based Water Quality Monitoring

Previous studies have used IoT sensors to monitor water quality parameters such as pH, temperature, and dissolved oxygen in real time. These systems successfully collect and transmit data but mainly rely on threshold-based alerts and do not use Artificial Intelligence for prediction or advanced analysis.

2. Machine Learning for Water Quality Assessment

Researchers have applied machine learning algorithms such as Support Vector Machine, Decision Tree, k-Nearest Neighbour, and Random Forest for water quality classification. Among these, Random Forest has shown higher accuracy and better performance. However, most studies use static datasets instead of real-time monitoring.

3. Anomaly Detection

Isolation Forest has been widely used to detect abnormal water quality conditions without requiring labelled data. It can identify pollution events efficiently with low computational cost, making it suitable for real-time environmental monitoring.

4. Explainable AI

Explainable AI helps users understand why AI makes a particular prediction. Existing methods improve trust but are computationally expensive. A lightweight rule-based explanation system provides simple and understandable reasons for pollution detection with lower processing time.

5. Digital Twin Technology

Digital Twin technology creates a virtual model of a real system for monitoring and analysis. Recent studies suggest that simulation-based Digital Twins are an effective way to validate smart water monitoring systems before deploying physical IoT sensors. This project combines AI, simulation, database management, and a web dashboard into a single integrated solution for smart lake monitoring.

III. METHODOLOGY

Step 1: Sensor Simulation

Instead of physical IoT sensors, this project uses a Digital Twin to simulate water quality data. The simulator generates pH, temperature, turbidity, and dissolved oxygen values every 5 seconds with realistic variations to represent actual lake conditions.

Step 2: Data Collection

The generated sensor readings are collected automatically and stored with timestamps. Each reading is immediately sent to the AI processing module for analysis.

Step 3: Data Storage

The collected data is stored in an SQLite database, which provides reliable and structured storage. The database can be easily replaced with a cloud database in future deployments.

Step 4: AI Processing

The system uses Random Forest to classify water quality as Good, Moderate, or Poor and Isolation Forest to detect abnormal conditions. A rule-based Explainable AI module provides simple reasons for the detected pollution.

Step 5: Prediction and Analysis

The system calculates a lake health score and analyzes recent trends to predict possible pollution risks. If water quality parameters indicate increasing pollution, an early warning alert is generated to support timely action and effective lake management.

ARCHITECTURE DIAGRAM

SLIM-AI Lite System Architecture

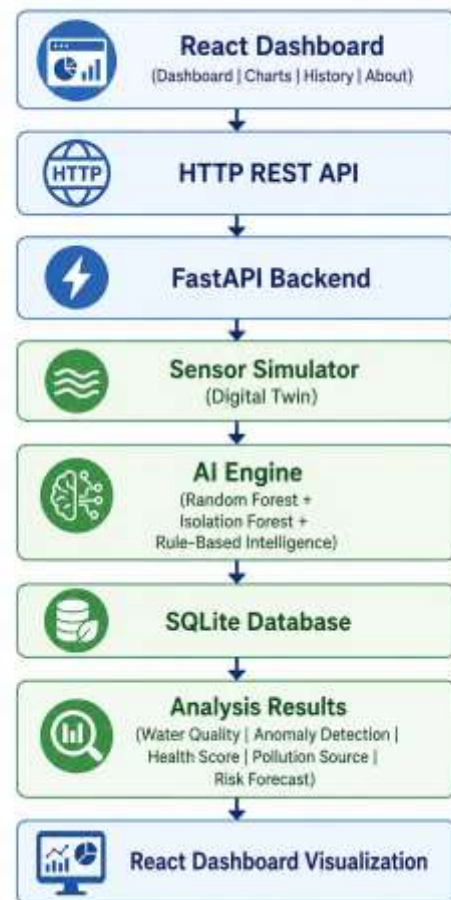


Fig 1: System Architecture

The architecture follows a layered design. The React Frontend at the top communicates via HTTP REST to the FastAPI server. The server hosts the Sensor Simulator, SQLite database, ML models (Random Forest and Isolation Forest), and the Rule-based Engine. The simulation loop continuously generates data, which flows through the AI pipeline, gets stored, and is served via the API to the frontend every 5 seconds.

IV. APPLIED TECHNOLOGY

To enable real-time water quality monitoring, intelligent environmental analysis, anomaly detection, and interactive visualization, the SLIM-AI Lite project integrates Artificial Intelligence, Machine Learning, Digital Twin technology, and modern web development frameworks. These technologies

collectively provide an efficient, scalable, and cost-effective smart lake monitoring system.

4.1 Python

Goal: Primary programming language for AI development, simulation, and backend implementation.

Use in Project: Python is used to develop the Digital Twin sensor simulator, implement Machine Learning models, perform data processing, interact with the SQLite database, and build the FastAPI backend.

Python offers extensive support for scientific computing and machine learning libraries, enabling rapid development, easy integration, and efficient execution of the complete environmental monitoring pipeline.

4.2 Visual Studio Code (VS Code)

Goal: Integrated Development Environment (IDE) for software development.

Use in Project: Used for writing, debugging, and managing the backend Python code, frontend React application, and project files.

VS Code provides intelligent code completion, debugging support, integrated terminal, and extension support, improving developer productivity throughout the project.

4.3 Scikit-learn

Goal: Machine Learning library for classification and anomaly detection.

Use in Project: Implements the Random Forest Classifier for water quality prediction and Isolation Forest for anomaly detection.

Scikit-learn provides reliable machine learning algorithms with high accuracy, low computational complexity, and easy integration into real-time applications.

4.4 FastAPI

Goal: High-performance Python web framework for REST API development.

Use in Project: Provides REST APIs that connect the backend AI engine, database, and frontend dashboard.

FastAPI delivers fast execution, automatic API documentation, asynchronous processing, and efficient communication between all project components.

4.5 React.js

Goal: Frontend JavaScript library for interactive user interfaces.

Use in Project: Develops the real-time monitoring dashboard displaying sensor readings, water quality status, anomaly alerts, health scores, and graphical trends.

React.js enables responsive, dynamic, and user-friendly visualization with real-time data updates.

4.6 SQLite

Goal: Lightweight relational database management system.

Use in Project: Stores sensor readings, AI prediction results, anomaly reports, and historical environmental data.

SQLite provides reliable local storage without requiring additional database servers, making it ideal for prototype and edge-based deployments.

4.7 Digital Twin Technology

Goal: Simulate real-world lake monitoring without physical hardware.

Use in Project: A Python-based sensor simulator generates realistic water quality parameters including pH, temperature, turbidity, and dissolved oxygen.

Digital Twin technology enables complete validation of AI models before deployment with actual IoT sensors, reducing development cost and complexity.

4.8 React Chart.js

Goal: Interactive graphical visualization of environmental data.

Use in Project: Displays real-time trends of pH, turbidity, temperature, and dissolved oxygen through dynamic line charts.

Chart.js helps users easily analyze water quality changes, identify pollution events, and monitor long-term environmental trends.

4.9 SQLite Database

Goal: Store and manage sensor and analysis data.

Use in Project: Maintains historical sensor readings, AI predictions, health scores, pollution source information, and anomaly detection results.

The database supports efficient querying, historical analysis, and future migration to cloud databases such as PostgreSQL or Firebase.

4.10 Windows/Linux Operating System

Goal: Development and execution environment.

Use in Project: Supports backend development, AI model execution, React frontend, and database management.

Both operating systems provide stable, cross-platform compatibility and reliable execution for AI-based environmental monitoring applications.

V. RESULTS AND DISCUSSION

5.1 Normal Operation Good Water Quality



Figure5.1: Dashboard during normal operation. All KPI cards show green/cyan status.

5.2 Anomaly Detection Warning State



Figure 5.2: Dashboard showing a detected pH anomaly

5.3 Simulated Pollution Event Poor Quality



Figure 5.3: Dashboard after triggering a simulated pH acidic pollution event.

5.4 Dashboard showing MODERATE water quality

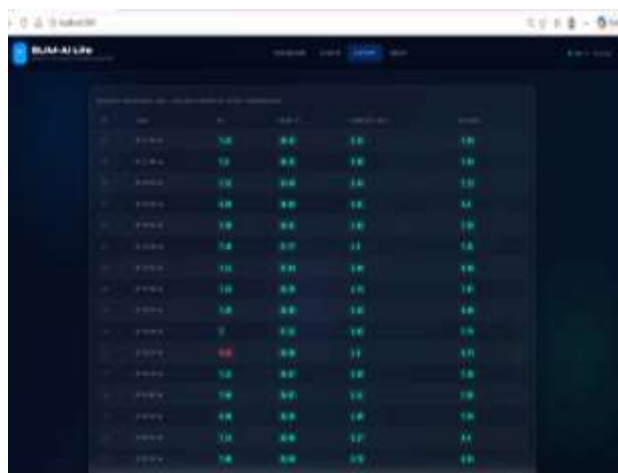


Figure 5.4: Dashboard showing MODERATE water quality (pH 5.21, acidic) with WARNING anomaly (score -0.0036).

5.5 Charts page showing a turbidity spike event

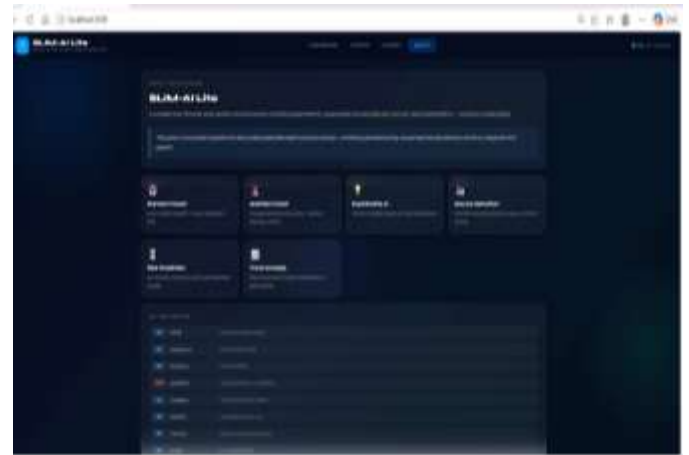


Figure 5.5: Charts page showing a turbidity spike event.

5.6 History page showing colour-coded recent readings log



Figure 5.6: History page showing colour-coded recent readings log.

5.7 Experiment Results

Test Case	Description	Expected Result	Actual Result	Status
TC1	Normal water quality detection	Classify water quality as good	Successfully Classified as good	Pass
TC2	Water quality classification	Classify water as good/moderate/poor	Correct classification achieved using Random Forest	Pass
TC3	Anomaly detection	Detect abnormal	Isolation forest	Pass

		water quality		
TC4	Pollution event simulation	Generate pollution alerts	System generate real -time pooltion anomalies	Pass
TC5	Lake health score calculation	Calculate lake health score	Accurate health score display	Pass
TC6	Pollution source identification	Identify probable pollution source	Correctly identified Industrial/organic/algae pollution	Pass
TC7	Water quality risk prediction	Predict future water quality risk	Successfull y generated 30 min every 5 seconds	Pass
TC8	Dashboard visualization	Display live sensor reading and charts	Real-time dashboard updated every 5 second	Pass
TC9	Historical data storage	Store sensor readings and ai analaysis	Data successfull y stored in SQLite database	Pass
TC1P	REST API communication	Fetch data through fastapi endpoints	API responded successfull y with real -time data	Pass

5.8 Discussion

The Random Forest classifier accurately classified water quality into Good, Moderate, and Poor categories with high confidence. The Isolation Forest successfully detected abnormal water quality conditions and classified them as Normal, Warning, or Critical. The Explainable AI (XAI) module clearly identified the parameters responsible for pollution and suggested possible pollution sources such as

industrial waste, domestic sewage, and algal blooms. The system also generated 30-minute risk predictions based on recent water quality trends, reducing false alerts by requiring continuous abnormal readings. The Digital Twin simulation enabled efficient testing of different pollution scenarios without physical sensors, demonstrating the effectiveness of the proposed AI-based lake monitoring system.

VI. CONCLUSION

This project, Smart Lake Rejuvenation with AI-Powered Environmental Monitoring (SLIM-AI Lite), successfully developed an AI-based system for continuous lake water quality monitoring. The system uses Random Forest for water quality classification and Isolation Forest for anomaly detection. It also provides explainable results, identifies possible pollution sources, calculates a lake health score, and predicts future pollution risks. The Digital Twin approach enables realistic testing without physical IoT sensors, making the system cost-effective and easy to validate. The interactive dashboard displays real-time readings, alerts, and historical data for better decision-making. Overall, the proposed system provides an efficient, scalable, and intelligent solution for sustainable lake monitoring and can be enhanced in the future by integrating real IoT sensors and cloud technology.

REFERENCE

- [1] Bhardwaj, A., Gupta, R., & Kumar, S. (2020). IoT-Based Real-Time Water Quality Monitoring System for Smart Aquaculture. *International Journal of Engineering Research and Technology*, 9(8), 78–83.
- [2] Raju, K. L., & Reddy, P. S. (2021). Design and Implementation of a Multi-Parameter Water Quality Monitoring Device Using Raspberry Pi. *Journal of Sensors*, 2021, Article ID 5567921. <https://doi.org/10.1155/2021/5567921>
- [3] Singh, A., Patel, H., & Desai, M. (2021). Machine Learning Approaches for River Water Quality Classification in India. *Water Resources Management*, 35(4), 1345–1360. <https://doi.org/10.1007/s11269-021-02754-z>
- [4] Chen, Y., & Shi, P. (2022). Water Quality Index Prediction Using Gradient Boosting Machines: A Case Study of Taihu Lake. *Environmental Modelling and Software*, 148, 105282. <https://doi.org/10.1016/j.envsoft.2021.105282>
- [5] Kiliç, K. E., & Ozbay, E. (2021). Anomaly Detection in Urban Water Systems Using One-Class SVM. *Procedia Computer Science*, 185, 190–197. <https://doi.org/10.1016/j.procs.2021.05.020>
- [6] Zhang, L., Liu, J., & Wang, H. (2023). Isolation Forest-Based Unsupervised Anomaly Detection for Municipal Water Quality Data. *Water Science and Technology: Water Supply*, 23(1), 112–124. <https://doi.org/10.2166/ws.2022.432>

- [7] Gupta, S., & Mehta, K. (2023). Explainable Artificial Intelligence for Lake Water Quality Assessment: A LIME-Based Approach. *Environmental Informatics*, 42(3), 567–580.
- [8] Pedersen, M., & Larsen, J. (2022). Digital Twin of a Wastewater Treatment Plant for Scenario Simulation and Predictive Control. *Water Research*, 211, 118012. <https://doi.org/10.1016/j.watres.2021.118012>
- [9] Ramesh, A., Nair, S., & Das, S. (2023). Simulation-Based Validation of IoT-Enabled Lake Monitoring Systems Using Digital Twins. *IEEE Sensors Journal*, 23(15), 17210–17220. <https://doi.org/10.1109/JSEN.2023.3275621>
- [10] Sharma, D., & Patel, R. (2024). A Review of Smart Water Quality Monitoring: From Sensors to Machine Learning and Digital Twins. *Sustainable Computing: Informatics and Systems*, 41, 100912. <https://doi.org/10.1016/j.suscom.2023.100912>
- [11] Liu, T., Xia, J., & Chen, W. (2022). Real-Time Water Quality Monitoring and Prediction Using IoT and Deep Learning for Smart Lake Management. *IEEE Internet of Things Journal*, 9(18), 17842–17856.
- [12] Bureau of Indian Standards. (2012). IS 10500: Indian Standard Drinking Water - Specification (Second Revision). BIS, New Delhi.
- [13] World Health Organization. (2022). Guidelines for Drinking-Water Quality: Fourth Edition Incorporating the First and Second Addenda. WHO Press, Geneva.